R-Lab 01

Moushreeta Debroy

2022-09-20

AIM: We draw a random sample of size 15 from the ToothGrowth dataset using without replacement procedure. Our objective is the following: i)Estimate the average cell length responsible for tooth growth in pig ii)Obtain the sampling distribution of the estimate and represent it by a histogram iii)TO find the mean and sd of the following iv)to increase sample size to see how it affects precision v)to find the standard error

ABOUT THE DATASET: The dataset which we'll be using for this study is the ToothGrowth. It is available in Base R. We are basically going to work with the length column in the dataset.

INTRODUCTION Here the sampling technique used simple random sampling without replacement. Sampling Distribution is the probability distribution followed by the test statistic of the sample.

ANALYSIS:

Now we perform some exploratory data analysis and get a sneakpeak into the dataset:

```
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

We look at the summary of the data as well

```
summary(ToothGrowth)
```

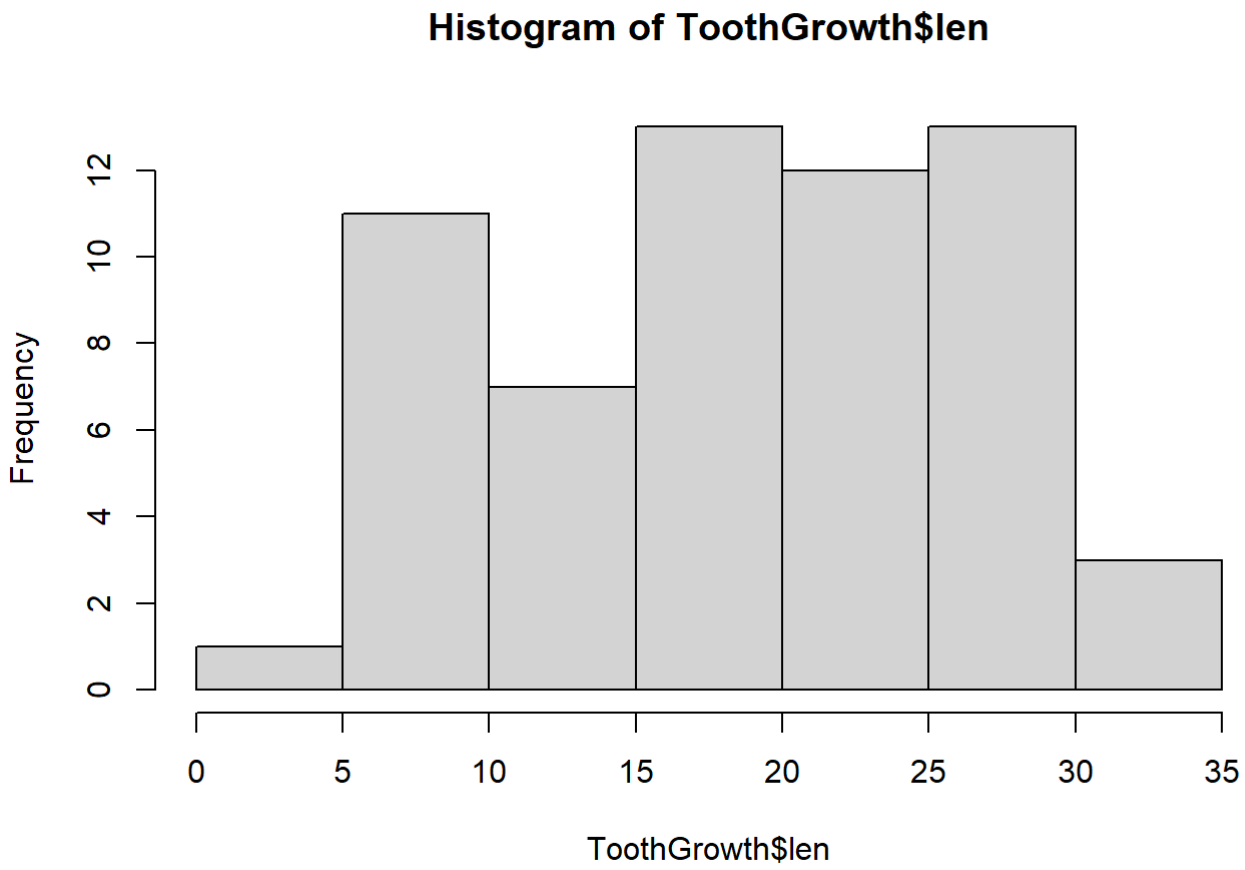
```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

The max value of the len variable is 33.90 and the min value of len is 4.20

```
tail(ToothGrowth)
```

```
##     len supp dose
## 55 24.8   OJ    2
## 56 30.9   OJ    2
## 57 26.4   OJ    2
## 58 27.3   OJ    2
## 59 29.4   OJ    2
## 60 23.0   OJ    2
```

```
hist(ToothGrowth$len)
```




```
mean(ToothGrowth$len)
```

```
## [1] 18.81333
```

```
sd(ToothGrowth$len)
```

```
## [1] 7.649315
```

Standard Devion closer to zero means that the data points are closer to mean and that the variance is lesser. In this case we have a sd of 7.65, meaning we have a considerabke variance

```
population=ToothGrowth$len
samplesize=15
```

Now we chose a sample size of 15 observations using simple random sampling without replacement procedure

```
sample=sample(population, samplesize, replace=F)
```

Now we perform exploratory data analysis on the sample we chose from the population

```
mean(sample)
```

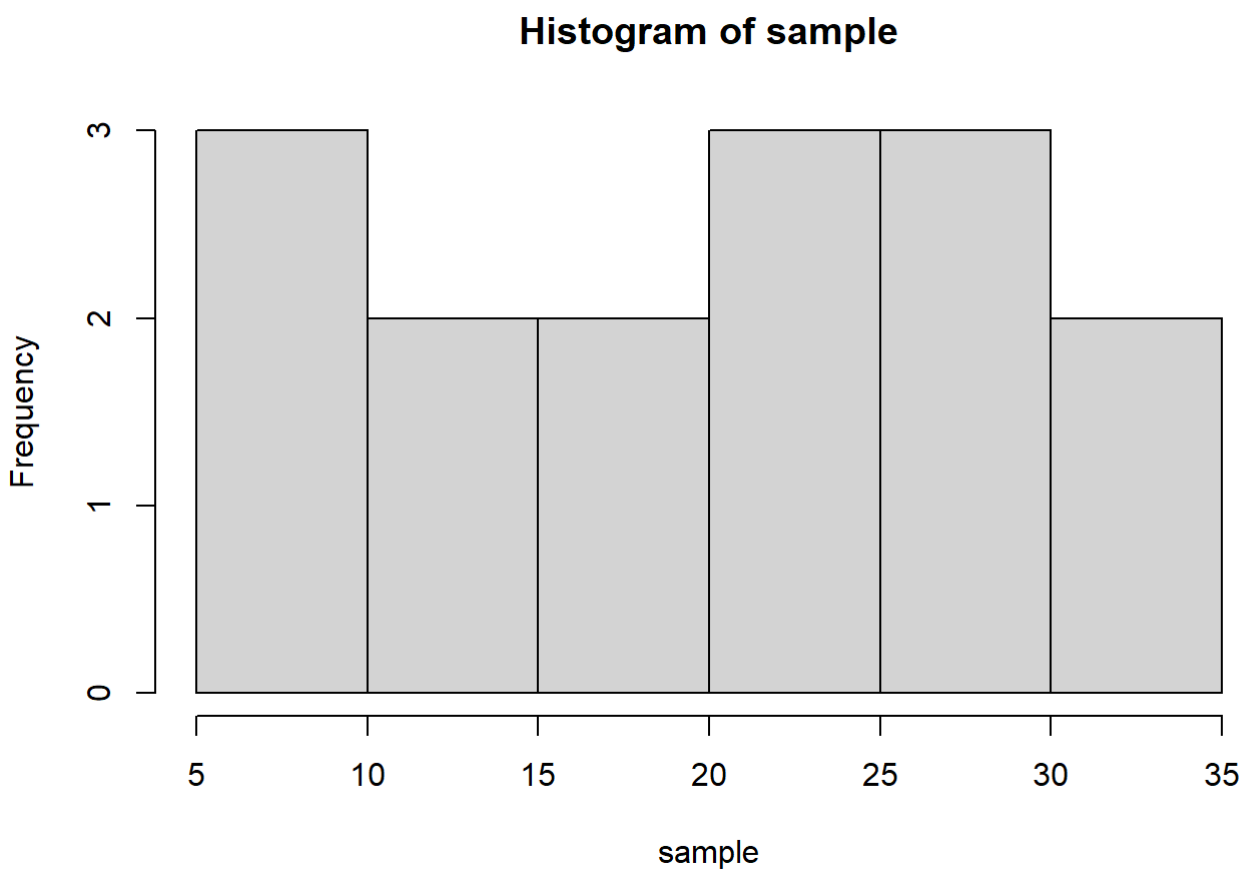
```
## [1] 19.82667
```

```
sd(sample)
```

```
## [1] 9.037103
```

Now we look at the distribution of the sample using a distribution

```
hist(sample)
```

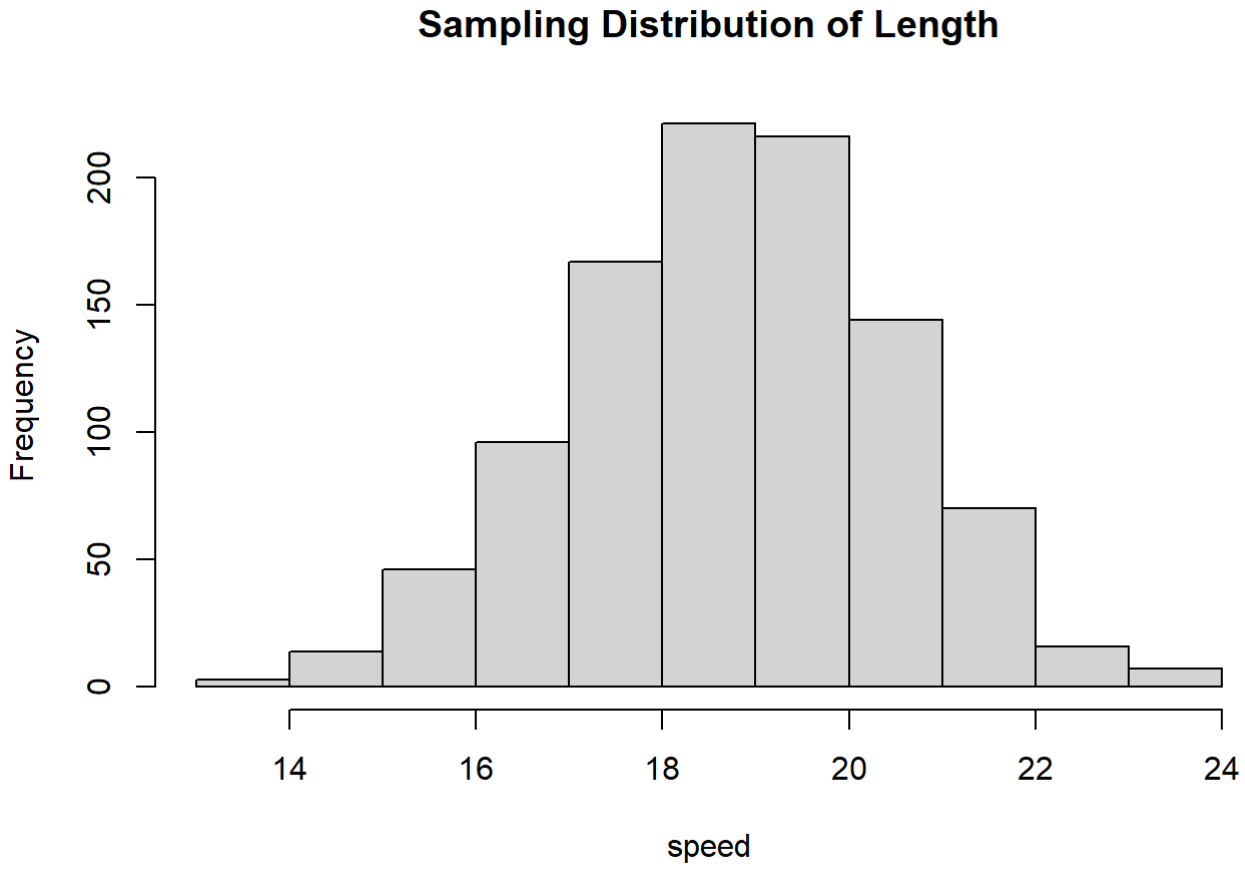


We replicate the sample 1000 times to get a better fit of the distribution. To do so we use the replicate() function repeatedly.

```
samp_dist=replicate(1000,mean(sample(ToothGrowth$len,15,replace=FALSE)))
```

Now to we look at the distribution again:

```
hist(samp_dist,main="Sampling Distribution of Length", xlab="speed")
```




```
var(samp_dist)
```

```
## [1] 2.993725
```

Now we'll increase the sample size to 30 and then calculate the variance.

```
samp_dist=replicate(1000,mean(sample(ToothGrowth$len,30,replace=FALSE)))
samp_dist
```

```
##    [1] 17.75667 17.45333 19.78000 20.32000 19.11667 18.51667 19.37000 18.17333
##    [9] 18.79667 19.48667 18.27667 18.43333 17.99333 20.58333 17.99000 20.05667
##   [17] 17.65000 19.33000 18.11667 18.12667 20.48333 20.93667 19.07000 19.09000
##   [25] 17.35333 18.36000 18.43667 18.58000 17.65667 19.25000 18.85667 18.57667
##   [33] 18.47000 18.58333 18.13000 20.97667 19.67333 19.33667 19.82667 19.64000
##   [41] 20.46667 18.32333 19.98000 20.09667 20.07667 19.66000 20.30000 19.44667
##   [49] 17.42333 18.57000 18.75000 19.58667 20.23000 18.76667 19.14333 18.98667
##   [57] 18.55667 17.66333 16.87333 19.58000 19.91333 19.37667 18.07667 17.95667
##   [65] 19.20667 16.98333 16.77000 18.41667 18.85000 18.46333 19.77000 20.19000
##   [73] 19.53000 20.45667 19.03667 19.55333 19.64667 19.41333 17.88000 18.30000
##   [81] 18.83333 17.90333 18.20667 18.58000 20.52667 17.22000 20.31000 18.64333
##   [89] 17.55000 21.53333 17.04667 18.07333 19.05667 19.01667 18.97667 18.34333
##   [97] 19.72667 17.74333 18.90667 18.59000 18.15667 18.06333 19.27000 18.96667
##  [105] 19.95000 20.10667 19.48000 18.25000 19.72667 16.96333 20.13000 20.22333
##  [113] 18.88000 17.99667 18.03667 17.85333 18.57000 19.08333 19.92333 17.41667
##  [121] 17.63333 18.48000 16.01000 17.80000 19.44333 18.35667 20.07000 18.35000
##  [129] 17.26000 19.31000 17.89000 19.02000 18.36667 19.41667 18.43333 18.80667
##  [137] 20.26000 18.29000 18.24667 18.89333 17.83000 17.51667 17.86333 17.00000
##  [145] 19.13667 19.39000 17.48333 19.13667 20.45000 17.80333 17.29333 18.72667
##  [153] 17.98000 19.59000 18.23667 18.70333 19.01000 19.03667 19.05667 19.07000
##  [161] 18.75000 18.83000 18.55667 17.88667 17.31000 19.20000 20.78000 17.37333
##  [169] 19.73000 18.75333 19.55667 18.70333 19.01667 19.67333 16.78333 18.45000
##  [177] 18.87333 18.64667 20.16667 19.79000 16.43333 18.36000 17.76000 19.36333
##  [185] 17.04333 19.11333 18.18667 17.32000 17.54667 19.22667 18.04000 18.55333
##  [193] 18.40000 18.84667 18.29000 19.98333 15.89333 17.15000 18.96667 18.39000
##  [201] 19.14333 18.91333 18.17667 18.99000 18.81000 19.07667 20.34667 19.31000
##  [209] 18.35667 20.06333 18.50000 17.71000 18.73667 18.56667 18.79667 18.27333
##  [217] 19.04000 19.65667 16.90000 17.31667 18.95000 20.72000 19.01000 19.31333
##  [225] 17.16667 19.85000 20.93333 19.22333 18.14333 19.94333 20.03000 18.61333
##  [233] 19.23000 18.13000 16.74000 18.46333 19.70000 18.69000 18.81333 18.89000
##  [241] 20.28667 19.27333 19.71333 18.61667 19.31333 19.21000 17.03667 19.58000
##  [249] 17.75667 18.92667 20.76667 17.38333 18.53000 19.51667 17.07667 17.70667
##  [257] 17.90667 20.11000 18.18667 18.86667 19.00333 18.76000 19.43333 20.15667
##  [265] 19.98000 19.19000 19.26667 20.00667 18.46333 18.28000 19.00667 17.62000
##  [273] 19.22000 18.24000 19.89667 17.19667 18.85333 16.60333 17.61333 17.97667
##  [281] 18.59667 17.45667 18.70333 19.80333 19.28667 18.92333 19.42000 18.86000
##  [289] 18.57667 18.69667 18.50333 18.54333 18.81333 19.67000 17.57000 18.55667
##  [297] 18.44000 20.31667 19.46333 19.74333 20.83667 18.44667 19.12667 18.36000
##  [305] 19.98667 18.42000 18.14000 19.76000 18.80000 18.91667 17.99000 18.85000
##  [313] 18.14000 19.08000 18.01000 19.09000 19.07333 17.90333 21.61667 17.72000
##  [321] 17.94000 19.41333 18.56000 18.74000 18.85000 17.65333 19.91333 17.05000
##  [329] 18.68333 16.92333 19.42000 17.70667 18.32333 20.31333 17.51333 19.59667
##  [337] 19.11000 18.32000 17.97667 19.18000 18.88333 18.11667 19.23667 20.12333
##  [345] 18.51333 19.82333 20.36333 18.16000 17.67000 16.97333 18.34667 18.72667
##  [353] 18.89000 19.37000 19.65667 18.23333 20.63667 18.92333 19.90333 18.21000
##  [361] 19.34000 21.69000 17.66333 18.13333 18.25000 18.26667 18.32333 18.66333
##  [369] 19.78667 18.86333 19.62333 19.19667 18.82333 16.88667 18.98333 19.11667
##  [377] 17.95333 18.08667 19.79333 18.76333 18.06667 18.14667 18.69667 19.77000
##  [385] 17.54333 19.37000 19.27667 18.50333 19.16000 16.16000 18.39667 17.79667
##  [393] 18.59333 17.83667 17.85000 17.03333 21.21333 18.61333 18.71000 17.34333
##  [401] 17.91000 18.31667 17.69667 18.35333 17.01333 18.94333 19.46667 20.50667
##  [409] 19.04333 18.76000 18.06667 20.83333 17.70667 20.36333 19.12000 15.99000
##  [417] 20.37000 17.13000 17.77000 19.63333 18.19000 18.30667 19.87000 17.30333
##  [425] 19.02000 19.24000 18.95000 19.31333 17.94333 18.95000 19.99000 19.63333
##  [433] 18.34000 17.17667 18.52000 18.55333 18.77667 19.88667 19.28667 18.84667
```

```
##  [441] 19.01000 19.93667 18.19000 18.76667 19.15000 17.38000 19.00667 17.72000
##  [449] 19.69333 20.14667 17.45333 17.99667 20.14667 19.01000 20.70667 19.60667
##  [457] 19.47000 18.99333 19.65667 18.75333 18.74000 19.95667 18.53333 18.66000
##  [465] 20.53333 20.06333 18.54667 19.77333 17.93667 18.82000 18.73000 19.66000
##  [473] 20.44667 18.72667 20.20667 19.50000 19.87000 16.49000 19.29667 19.77000
##  [481] 20.32667 20.30000 20.26000 17.21667 18.85667 18.19667 18.14333 17.76000
##  [489] 19.63000 17.52333 19.51667 19.49667 19.52667 18.86000 18.66333 18.33667
##  [497] 17.58333 19.72333 17.20333 17.40333 18.77333 18.88667 17.33000 17.98000
##  [505] 18.66000 18.23667 19.73667 19.11000 19.68333 19.87000 18.10000 19.21000
##  [513] 17.59333 18.33667 18.11333 19.13667 19.63333 20.10333 18.94333 18.93333
##  [521] 18.74667 18.89333 18.93000 18.09000 19.92333 20.03333 19.98333 18.72667
##  [529] 18.38000 18.22000 17.58000 20.12667 20.48333 20.60667 18.93667 17.57333
##  [537] 19.84000 16.86000 17.60000 20.08333 18.39333 18.08667 18.99667 17.96667
##  [545] 16.91000 19.29333 20.20667 18.06000 19.12667 17.94667 16.84667 19.65000
##  [553] 18.53333 18.46000 16.44000 18.90000 20.18333 18.92000 19.48000 18.67667
##  [561] 18.94667 16.65333 19.08000 19.13000 17.70000 19.52333 19.72333 17.43333
##  [569] 18.76333 17.92667 17.43667 19.73333 19.05667 19.91000 19.52667 17.93000
##  [577] 17.81333 17.78667 17.84667 18.74667 20.68000 18.85667 19.28667 18.49333
##  [585] 18.65333 20.35667 17.84333 18.91667 17.77000 19.30667 17.03667 18.50333
##  [593] 18.15000 19.50667 18.71667 17.33667 17.37333 19.16667 19.90000 19.47667
##  [601] 17.94333 19.16333 16.04000 17.39667 17.13333 18.12667 19.42000 18.12333
##  [609] 19.17667 18.09000 19.86333 17.59000 19.06667 17.59333 17.00667 19.77333
##  [617] 18.91667 19.45667 17.96333 17.71000 19.09000 16.84333 21.55000 16.99000
##  [625] 20.02667 17.41000 18.13000 20.19667 19.81333 19.67333 18.38333 20.41000
##  [633] 18.93000 19.57333 18.61333 18.61667 20.21000 18.40667 17.30667 19.36333
##  [641] 18.60000 18.81000 18.50333 18.70000 18.18333 19.06000 17.22667 18.61333
##  [649] 18.07333 18.60000 18.95333 21.10333 19.95667 18.36000 18.85333 18.27000
##  [657] 18.63333 20.23333 19.46667 16.76333 18.96000 19.31667 19.37333 19.56000
##  [665] 18.83333 19.88333 16.74000 19.71667 20.02000 19.09667 19.87333 16.93333
##  [673] 20.60000 17.54000 19.02000 19.30000 19.82000 21.46667 18.11000 19.88667
##  [681] 19.87667 18.19667 17.35333 18.76667 18.91667 17.44667 19.72333 16.77333
##  [689] 19.54000 18.41333 19.39000 18.22000 19.44000 18.56667 18.53000 17.48333
##  [697] 19.23000 18.07667 19.44333 19.30667 18.51667 18.93667 19.46000 16.75000
##  [705] 19.62333 18.76000 19.97667 19.59667 19.08667 19.75333 18.34667 18.50333
##  [713] 18.14000 19.98000 17.88000 19.61667 18.53000 17.64667 17.26000 19.25000
##  [721] 18.56667 17.71333 19.01333 19.92000 18.22333 19.13000 20.32000 20.21667
##  [729] 19.51333 17.24667 17.10667 20.86000 18.01333 19.34333 17.42000 19.38000
##  [737] 19.49667 17.84333 20.60333 19.85667 18.43000 18.27667 18.33333 19.51667
##  [745] 19.02000 18.85000 16.99000 19.65333 18.88333 20.49000 18.43667 16.87000
##  [753] 19.17333 19.94333 19.37667 19.21333 20.39000 19.53667 20.56000 21.52000
##  [761] 19.91667 20.82000 19.92000 19.45667 17.39000 18.54667 19.47333 18.54667
##  [769] 18.67333 19.02333 17.37667 18.66333 18.75000 17.82333 17.62667 17.74667
##  [777] 19.25000 19.25333 19.72000 19.72667 18.57000 20.56333 20.13000 19.79667
##  [785] 19.45000 19.42667 18.83000 17.50667 21.15667 19.98333 19.46000 20.27667
##  [793] 19.24000 20.57667 18.18333 18.70667 20.47000 17.25667 19.30000 18.62667
##  [801] 17.01667 18.25667 18.13667 18.80333 18.74333 18.49333 19.02333 19.37000
##  [809] 18.79000 19.79333 18.22333 19.35667 18.09000 20.06000 18.14333 20.39333
##  [817] 17.79333 19.52000 18.45667 19.14667 18.34667 19.61333 18.85667 16.68333
##  [825] 18.88667 20.27333 17.64000 18.13333 18.30333 19.27667 18.51667 18.15667
##  [833] 18.70000 19.23667 18.54000 20.47667 19.52333 18.03000 20.53333 17.79000
##  [841] 19.24667 18.10333 18.98000 19.27667 20.44667 19.23667 17.43333 16.84667
##  [849] 18.57667 19.80000 19.46000 17.88333 17.57667 19.77000 18.97333 17.97333
##  [857] 18.84000 17.44333 20.49333 19.87667 18.70667 19.41000 18.56333 18.57333
##  [865] 19.37667 19.18333 17.34333 17.05333 20.00333 19.82667 17.15000 16.71667
##  [873] 18.60000 20.23333 18.67000 19.66000 19.27333 18.39667 18.66667 20.81667
##  [881] 18.15000 20.80667 16.24000 18.35333 20.88667 18.48000 18.65000 17.66333
```

```
##  [889] 21.45000 19.14000 17.10667 17.65667 18.28333 18.87000 19.01667 19.21333
##  [897] 18.89667 17.17333 17.48000 19.39000 19.33333 20.46667 19.01333 17.66333
##  [905] 18.76000 17.90667 19.61000 18.91333 18.82333 18.86333 16.25333 20.02000
##  [913] 18.52000 19.97667 18.61000 18.93667 18.95667 18.24667 17.04667 19.27667
##  [921] 19.04000 18.08333 19.13333 18.00000 19.57667 19.62000 19.29000 18.73667
##  [929] 19.59000 18.71667 19.10333 18.94000 20.17000 18.55667 17.03333 18.15000
##  [937] 19.02667 19.52333 20.62667 17.95000 19.40000 18.03000 19.06333 19.16667
##  [945] 17.71667 19.81667 17.22000 17.89000 19.57667 18.47333 20.85333 19.98000
##  [953] 19.40667 17.83000 19.99000 18.37333 18.91333 19.04000 19.88000 19.42000
##  [961] 17.14333 19.05667 17.13667 16.88333 19.09667 19.38333 17.83333 20.13333
##  [969] 18.20333 19.57667 19.27333 18.71000 18.62333 17.05000 20.09667 20.40000
##  [977] 18.16333 18.42000 19.01000 17.65000 18.94667 18.00000 18.44667 19.58667
##  [985] 18.04667 18.75667 20.17667 18.07667 16.99667 18.19000 19.22000 18.53667
##  [993] 18.05333 20.10667 19.86000 19.78000 18.99667 17.94667 20.29333 18.43667
```

Now we check the variance once again.

```
var(samp_dist)
```

```
## [1] 1.017297
```

Here, we see that as the sample size increases the variance decreases which in turn increases the precision.

```
samplesize=25
print(sd(samp_dist)/sqrt(samplesize))
```

```
## [1] 0.2017222
```

CONCLUSION: Here we see that increasing the sample size increases the precision.The standard error of the sample is 0.1941487.